StyleNeRF: A Style-based 3D-Aware Generator for High-resolution Image Synthesis

Abstract

We propose StyleNeRF, a 3D-aware generative model for photo-realistic high-resolution image synthesis with high multi-view consistency, which can be trained on unstructured 2D images. Existing approaches either cannot synthesize high-resolution images with fine details or yield noticeable 3D-inconsistent artifacts. In addition, many of them lack control over style attributes and explicit 3D camera poses. StyleNeRF integrates the neural radiance field (NeRF) into a style-based generator to tackle the aforementioned challenges, i.e., improving rendering efficiency and 3D consistency for high-resolution image generation. We perform volume rendering only to produce a low-resolution feature map and progressively apply upsampling in 2D to address the first issue. To mitigate the inconsistencies caused by 2D upsampling, we propose multiple designs, including a better upsampler and a new regularization loss. With these designs, StyleNeRF can synthesize high-resolution images at interactive rates while preserving 3D consistency at high quality. StyleNeRF also enables control of camera poses and different levels of styles, which can generalize to unseen views. It also supports challenging tasks, including zoom-in and-out, style mixing, inversion, and semantic editing.